Reaching the End of Unbiasedness: Uncovering Implicit Limitations of Click-Based Learning to Rank

Harrie Oosterhuis

July 12, 2022

Radboud University
harrie.oosterhuis@ru.nl
https://twitter.com/HarrieOos

Main Question

How far can the prevalent approach in the **Unbiased Learning-to-Rank** field take us?

In Other Words

What are the limits of the field's current approach to **Unbiased Learning-to-Rank**?

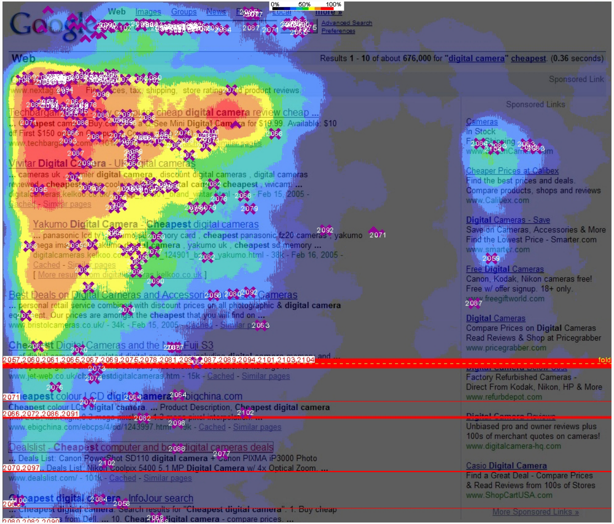
Introduction:
Click-Based Learning-to-Rank

Click-Based Learning-to-Rank:

- Optimization of **ranking models** w.r.t. ranking metrics for **search** or **recommendation** based on **user-interactions** (i.e. clicks).
- Encapsulates unbiased / counterfactual / offline / online LTR.

Core Problem:

- **Mismatch between click-probabilities and relevances.**
- **Other factors** beside relevance affect CTR as well.



Source: http://www.mediative.com/

Existing Click-Based LTR Approach

Vardasbi et al. (2020) **assume** the click model, with doc. d, query q, rank k, relevance $R_{d|q}$, bias parameters α_k and β_k per rank:

$$P(C = 1 \mid d, k, q) = \alpha_k R_{d|q} + \beta_k.$$

A. Vardasbi, H. Oosterhuis, and M. de Rijke. When inverse propensity scoring does not work: Affine corrections for unbiased learning to rank. CIKM 2020.

Vardasbi et al. (2020) **assume** the click model, with doc. d, query q, rank k, relevance $R_{d|q}$, bias parameters α_k and β_k per rank:

$$P(C = 1 \mid d, k, q) = \alpha_k R_{d|q} + \beta_k.$$

They propose an estimator based on the **inverse transformation**:

$$\hat{R}_{d|q} = \frac{1}{N} \sum_{i=1}^{N} \frac{c_i(d) - \hat{\beta}_{k_i(d)}}{\hat{\alpha}_{k_i(d)}}.$$

Straightforwardly **unbiased** if bias is correctly estimated:

$$\left(\hat{\alpha} = \alpha \wedge \hat{\beta} = \beta\right) \longrightarrow \mathbb{E}\left[\hat{R}_{d|q}\right] = R_{d|q}.$$

A. Vardasbi, H. Oosterhuis, and M. de Rijke. When inverse propensity scoring does not work: Affine corrections for unbiased learning to rank. CIKM 2020.

A **basic click model** estimates by **optimizing parameters** of a predictive model:

$$\hat{P}(C = 1 \mid d, k, q) = \hat{\alpha}_k \hat{R}_{d|q} + \hat{\beta}_k.$$

A **basic click model** estimates by **optimizing parameters** of a predictive model:

$$\hat{P}(C = 1 \mid d, k, q) = \hat{\alpha}_k \hat{R}_{d|q} + \hat{\beta}_k.$$

The values of R, α and β are inferred by **minimizing** a **loss** function:

$$\hat{\mathcal{L}} = \frac{-1}{N} \sum_{i=1}^{N} c_i(d) \log\left(\hat{\alpha}_{k_i(d)} \hat{R}_{d|q} + \hat{\beta}_{k_i(d)}\right) + (1 - c_i(d)) \log\left(1 - \hat{\alpha}_{k_i(d)} \hat{R}_{d|q} - \hat{\beta}_{k_i(d)}\right).$$

A **basic click model** estimates by **optimizing parameters** of a predictive model:

$$\hat{P}(C = 1 \mid d, k, q) = \hat{\alpha}_k \hat{R}_{d|q} + \hat{\beta}_k.$$

The values of R, α and β are inferred by **minimizing** a **loss** function:

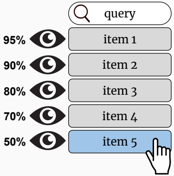
$$\hat{\mathcal{L}} = \frac{-1}{N} \sum_{i=1}^{N} c_i(d) \log\Big(\hat{\alpha}_{k_i(d)} \hat{R}_{d|q} + \hat{\beta}_{k_i(d)}\Big) + (1 - c_i(d)) \log\Big(1 - \hat{\alpha}_{k_i(d)} \hat{R}_{d|q} - \hat{\beta}_{k_i(d)}\Big).$$

The **expected** loss is the **cross-entropy loss** on click probabilities:

$$\mathbb{E}[-\hat{\mathcal{L}}] = P(C = 1 \mid d, q) \log \hat{P}(C = 1 \mid d, q) + P(C = 0 \mid d, q) \log \hat{P}(C = 0 \mid d, q).$$

Methodology:
Uncovering Limitations

Assumptions

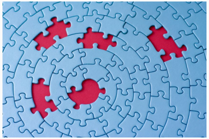


Assumed Click Model

Methods

Novel Counterfactual Estimator
or Click Modelling Method

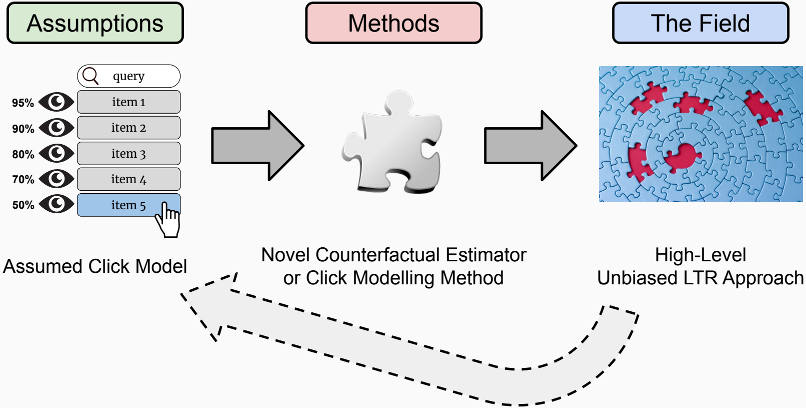
The Field



High-Level
Unbiased LTR Approach

Will this process eventually **solve all forms of bias**?

Methodology: Visualization



To recognize **limitations** we should go into the **other direction**.

Methodology:

- Find the most **generic definitions** for the **main approaches** in click-based LTR: **counterfactual estimation** and **click modelling**.

- From these definitions, **derive** under which **conditions** they cannot be **unbiased** or **consistent**:

$$\underbrace{\forall (d,q),\ \mathbb{E}\big[\hat{R}_{d|q}\big] = R_{d|q},}_{\text{unbiasedness}} \qquad \underbrace{\forall (d,q),\ \lim_{|\mathcal{D}_q|\to\infty} \hat{R}_{d|q} = R_{d|q}\,.}_{\text{consistency}}$$

Limitations of Click-Based Counterfactual Estimators

Display Context:

$x(d)$ contains **all information** about how d is **displayed**, but **no information** about **relevance** $R_{d|q}$.

Display Context:

$x(d)$ contains **all information** about how d is **displayed**,
but **no information** about **relevance** $R_{d|q}$.

Counterfactual Relevance Estimate:

Average over independently-sampled **interactions**, each **transformed** by function f:

$$\hat{R}_{d|q} = \frac{1}{N} \sum_{i=1}^{N} f(c_i(d), x_i(d)).$$

f only has **two relevant values** per context $x(d)$: $f(1, x(d))$ and $f(0, x(d))$.

Click-Based Counterfactual Estimators: Definitions

Display Context:

$x(d)$ contains **all information** about how d is **displayed**,
but **no information** about **relevance** $R_{d|q}$.

Counterfactual Relevance Estimate:

Average over independently-sampled **interactions**, each **transformed** by function f:

$$\hat{R}_{d|q} = \frac{1}{N} \sum_{i=1}^{N} f(c_i(d), x_i(d)).$$

f only has **two relevant values** per context $x(d)$: $f(1, x(d))$ and $f(0, x(d))$.

This definition captures **virtually all** existing counterfactual LTR methods.

Derivation in a **nutshell**:

$$\mathbb{E}_{c,x}\left[\hat{R}_{d|q}\right] = R_{d|q} \longleftrightarrow \lim_{|\mathcal{D}_q|\to\infty} \hat{R}_{d|q} = R_{d|q} \longleftrightarrow \mathbb{E}_{c,x}[f(c(d), x(d))] = R_{d|q},$$

Derivation in a **nutshell**:

$$\mathbb{E}_{c,x}\left[\hat{R}_{d|q}\right] = R_{d|q} \longleftrightarrow \lim_{|\mathcal{D}_q|\to\infty} \hat{R}_{d|q} = R_{d|q} \longleftrightarrow \mathbb{E}_{c,x}[f(c(d), x(d))] = R_{d|q},$$

the **expected value** can be rewritten to:

$$\begin{aligned}
&\mathbb{E}_{c,x}[f(c(d), x(d)) \mid q] \\
&\qquad = \mathbb{E}_{x}[P(C = 1 \mid d, x, q) f(1, x(d)) + (1 - P(C = 1 \mid d, x, q)) f(0, x(d)) \mid q] \\
&\qquad = \mathbb{E}_{x}[P(C = 1 \mid d, x, q)(f(1, x(d)) - f(0, x(d))) + f(0, x(d)) \mid q].
\end{aligned}$$

Derivation in a **nutshell**:

$$\mathbb{E}_{c,x}\left[\hat{R}_{d|q}\right] = R_{d|q} \longleftrightarrow \lim_{|\mathcal{D}_q|\to\infty} \hat{R}_{d|q} = R_{d|q} \longleftrightarrow \mathbb{E}_{c,x}[f(c(d), x(d))] = R_{d|q},$$

the **expected value** can be rewritten to:

$$\begin{aligned}
&\mathbb{E}_{c,x}[f(c(d), x(d)) \mid q] \\
&\qquad = \mathbb{E}_{x}[P(C = 1 \mid d, x, q) f(1, x(d)) + (1 - P(C = 1 \mid d, x, q)) f(0, x(d)) \mid q] \\
&\qquad = \mathbb{E}_{x}[P(C = 1 \mid d, x, q)(f(1, x(d)) - f(0, x(d))) + f(0, x(d)) \mid q].
\end{aligned}$$

Thus unbiasedness or consistency are **only possible** when click probabilities are an **affine transformation** of relevance s.t:

$$R_{d|q} = \mathbb{E}_{x}\left[\frac{P(C = 1 \mid d, x, q) - \beta_{x(d)}}{\alpha_{x(d)}} \mid q\right].$$

From our **generic definition**, we derived a very **broad condition**:

$$R_{d|q} = \mathbb{E}_x\left[\frac{P(C = 1 \mid d, x, q) - \beta_{x(d)}}{\alpha_{x(d)}} \mid q\right],$$

but it is **hard to interpret** as it relies on the **distribution** over both x and c.

Remember: $x \sim$ logging policy, and $c \sim$ user behavior given x.

From our **generic definition**, we derived a very **broad condition**:

$$R_{d|q} = \mathbb{E}_x\left[\frac{P(C = 1 \mid d, x, q) - \beta_{x(d)}}{\alpha_{x(d)}} \mid q\right],$$

but it is **hard to interpret** as it relies on the **distribution** over both x and c.

Remember: $x \sim$ logging policy, and $c \sim$ user behavior given x.

If **assumptions** on the logging policy are **impossible**, or under a **deterministic** policy, we get the simpler condition:

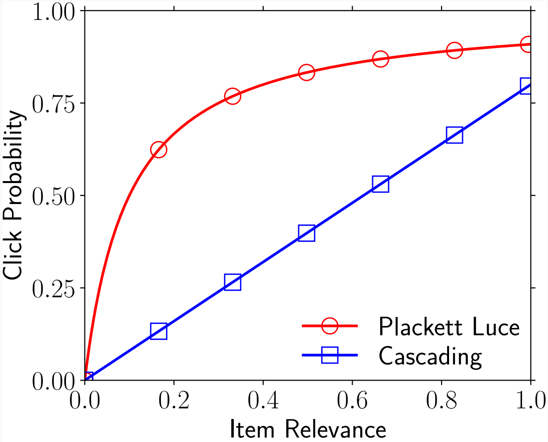
$$P(C = 1 \mid d, x, q) = \alpha_{x(d)} R_{d|q} + \beta_{x(d)}.$$

Click-Based Counterfactual Estimators: Visualization

Intuitive Interpretation:

- unbiasedness and consistency of counterfactual estimation are **only possible** under a **linear relation** between click probability and item relevance.

Reason is that click-based counterfactual estimates are **linear interpolations**.



Limitations of Click-Modelling

A click-model is defined by a **loss function** $\mathcal{L}(\hat{R}_q, \hat{Z}, \mathcal{D}_q)$ with as input:

- estimated **relevance** variables $\hat{R}_q$,
- estimated **latent** variables $\hat{Z}$,
- observed **data** $\mathcal{D}_q$.

The **relevance estimate** of a click-model are **any** values that **minimize** $\mathcal{L}$:

$$(\hat{R}_q^*, \hat{Z}^*) = \arg\min_{\hat{R}_q, \hat{Z}} \mathcal{L}(\hat{R}_q, \hat{Z}, \mathcal{D}_q).$$

A click-model is defined by a **loss function** $\mathcal{L}(\hat{R}_q, \hat{Z}, \mathcal{D}_q)$ with as input:

- estimated **relevance** variables $\hat{R}_q$,
- estimated **latent** variables $\hat{Z}$,
- observed **data** $\mathcal{D}_q$.

The **relevance estimate** of a click-model are **any** values that **minimize** $\mathcal{L}$:

$$(\hat{R}_q^*, \hat{Z}^*) = \arg\min_{\hat{R}_q, \hat{Z}} \mathcal{L}(\hat{R}_q, \hat{Z}, \mathcal{D}_q).$$

Captures both traditional **graphical models** and recent **deep-learning models**.

A click modelling method is consistent **iff** the **only relevance estimates** that **minimize** its **loss** are the **true relevances** as $|\mathcal{D}_q|$ tends to infinity:

$$\lim_{|\mathcal{D}_q|\rightarrow\infty} \big(R_q = \hat{R}_q \longleftrightarrow \min_{\hat{Z}} \mathcal{L}(\hat{R}_q, \hat{Z}, \mathcal{D}_q) = \min_{\hat{R}'_q, \hat{Z}} \mathcal{L}(\hat{R}'_q, \hat{Z}, \mathcal{D}_q)\big).$$

A click modelling method is consistent **iff** the **only relevance estimates** that **minimize** its **loss** are the **true relevances** as $|\mathcal{D}_q|$ tends to infinity:

$$\lim_{|\mathcal{D}_q|\to\infty} \big(R_q = \hat{R}_q \longleftrightarrow \min_{\hat{Z}} \mathcal{L}(\hat{R}_q, \hat{Z}, \mathcal{D}_q) = \min_{\hat{R}'_q, \hat{Z}} \mathcal{L}(\hat{R}'_q, \hat{Z}, \mathcal{D}_q)\big).$$

A click modelling method is unbiased **iff** the **expected value** of its optimal relevance estimates are equal to the **true relevances**:

$$\mathbb{E}_{\mathcal{D}_q}\big[\hat{R}^*_q\big] = R_q \longleftrightarrow \forall d,\ \mathbb{E}_{\mathcal{D}_q}\big[\hat{R}^*_{d|q}\big] = R_{d|q}.$$

Consistency depends **both** on the loss/model **and** how **data** is **gathered**.

Consistency depends **both** on the loss/model **and** how **data** is **gathered**.

	Ranking 1				Ranking 2			
Items	A	B	C	D	B	A	D	C
Click Prob.	0.90	0.64	0.40	0.05	0.8	0.72	0.20	0.10

Fitting a **rank-based position-bias model** to this scenario:

$$\hat{P}(C = 1 \mid d, k, q) = \hat{\alpha}_k \hat{R}_{d|q} \text{ with } \hat{\alpha}_1 = 1$$

Consistency depends **both** on the loss/model **and** how **data** is **gathered**.

	Ranking 1				Ranking 2			
Items	A	B	C	D	B	A	D	C
Click Prob.	0.90	0.64	0.40	0.05	0.8	0.72	0.20	0.10

Fitting a **rank-based position-bias model** to this scenario:

$$\hat{P}(C = 1 \mid d, k, q) = \hat{\alpha}_k \hat{R}_{d|q} \text{ with } \hat{\alpha}_1 = 1$$

Problem: Different relevance estimates **equally predictive**:

- $\hat{\alpha}_2 = 0.8,\ \hat{\alpha}_3 = 4 \cdot \hat{\alpha}_4,$
- $\hat{R}_A = 0.9,\ \hat{R}_B = 0.8$ and $\hat{R}_C = 2 \cdot \hat{R}_D.$

Conclusion

Conclusion

A theoretical investigation **implicit assumptions** in **Counterfactual Estimation** and **Click-Modelling** for **unbiased** and **consistent** learning-to-rank.

Ziniu Hu, Yang Wang, Qu Peng, and Hang Li. Unbiased LambdaMART: An Unbiased Pairwise Learning-to-Rank Algorithm. WWW 2019.

A theoretical investigation **implicit assumptions** in **Counterfactual Estimation** and **Click-Modelling** for **unbiased** and **consistent** learning-to-rank.

Main Findings

- Counterfactual estimation can only be unbiased when clicks follow **affine transformations** of relevance.

Ziniu Hu, Yang Wang, Qu Peng, and Hang Li. Unbiased LambdaMART: An Unbiased Pairwise Learning-to-Rank Algorithm. WWW 2019.

A theoretical investigation **implicit assumptions** in **Counterfactual Estimation** and **Click-Modelling** for **unbiased** and **consistent** learning-to-rank.

Main Findings

- Counterfactual estimation can only be unbiased when clicks follow **affine transformations** of relevance.
- Consistency of **click-modelling** methods depend on their **expected loss minima**; unclear if robust unbiasedness guarantees are possible.

Ziniu Hu, Yang Wang, Qu Peng, and Hang Li. Unbiased LambdaMART: An Unbiased Pairwise Learning-to-Rank Algorithm. WWW 2019.

A theoretical investigation **implicit assumptions** in **Counterfactual Estimation** and **Click-Modelling** for **unbiased** and **consistent** learning-to-rank.

Main Findings

- Counterfactual estimation can only be unbiased when clicks follow **affine transformations** of relevance.
- Consistency of **click-modelling** methods depend on their **expected loss minima**; unclear if robust unbiasedness guarantees are possible.
- **"Unbiased LambdaMART"** is **not** a sound debiasing method; illustrative example of why to be careful with assumptions.

Ziniu Hu, Yang Wang, Qu Peng, and Hang Li. Unbiased LambdaMART: An Unbiased Pairwise Learning-to-Rank Algorithm. WWW 2019.

Implications for the Future of the Field

For unbiasedness, we need **novel** methods that **differ** from our **generic** definitions:

- i.e. not **transforming individual** clicks,
- very **hard** to prove unbiasedness.

For unbiasedness, we need **novel** methods that **differ** from our **generic** definitions:

- i.e. not **transforming individual** clicks,
- very **hard** to prove unbiasedness.

Unbiasedness may **not** always be **possible**:

- we should **not** invariably **expect** nor require it of **future work**;

Implications for the Future of the Field

For unbiasedness, we need **novel** methods that **differ** from our **generic** definitions:

- i.e. not **transforming individual** clicks,
- very **hard** to prove unbiasedness.

Unbiasedness may **not** always be **possible**:

- we should **not** invariably **expect** nor require it of **future work**;
- unbiasedness may **not** be a **realistic long term goal**,
  field will likely **shift** to bias mitigation or **partial debiasing**;

For unbiasedness, we need **novel** methods that **differ** from our **generic** definitions:

- i.e. not **transforming individual** clicks,
- very **hard** to prove unbiasedness.

Unbiasedness may **not** always be **possible**:

- we should **not** invariably **expect** nor require it of **future work**;
- unbiasedness may **not** be a **realistic long term goal**,
  field will likely **shift** to bias mitigation or **partial debiasing**;
- good reason to **re-name** the field: **Click-Based LTR**.

Acknowledgments

Google

Radboud University

All content represents the opinion of the author, which is not necessarily shared or endorsed by their employers and/or sponsors.